43P

NsG-496

R. L. E. TECHNICAL REPORT 422

CODING COLOR PICTURES

URI F. GRONEMANN

FACILITY FORM 602

N64-27907 (ACCESSION NUMBER) — (THRU) 1

43 (PAGES) — (CODE) 15

CR-58170 (NASA CR OR TMX OR AD NUMBER) — (CATEGORY)

UNPUBLISHED PRELIMINARY DATA

TECHNICAL REPORT 422

JUNE 17, 1964

OTS PRICE

XEROX $ 4.60 ph

MICROFILM $ _____

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

RESEARCH LABORATORY OF ELECTRONICS

CAMBRIDGE, MASSACHUSETTS

RC# 10

The Research Laboratory of Electronics is an interdepartmental laboratory in which faculty members and graduate students from numerous academic departments conduct research.

The research reported in this document was made possible in part by support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the U.S. Army (Electronics Materiel Agency), the U.S. Navy (Office of Naval Research), and the U.S. Air Force (Office of Scientific Research) under Contract DA36-039-AMC-03200(E); and in part by Grant DA-SIG-36-039-61-G14; additional support was received from the National Science Foundation (Grant GP-2495), the National Institutes of Health (Grant MH-04737-04), and the National Aeronautics and Space Administration (Grant NsG-496).

Reproduction in whole or in part is permitted for any purpose of the United States Government.

CASE FILE COPY

4174

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

RESEARCH LABORATORY OF ELECTRONICS

Technical Report 422 June 17, 1964

CODING COLOR PICTURES

Uri F. Gronemann

Submitted to the Department of Electrical Engineering, M.I.T., February 1964, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

(Manuscript received January 15, 1964)

Abstract

27907

A computer-simulation study of efficient coding for color pictures has been undertaken. Two typical color transparencies were resolved into three primaries, sampled in a square array and recorded digitally on magnetic tape. The computer program transformed these data into luminance and chrominance quantities, performed certain parameter modifications, reconverted them into primary-color quantities and wrote them on an output tape. The parameters that were modified were the effective number of samples per picture and the number of quantum values each for the luminance and for the chrominance. The output tape was played back through the recorder-reproducer to produce images of the coded pictures on the face of the cathode-ray tube, which were photographed through appropriate filters on color film. The resultant transparencies were later viewed and compared by a number of observers to determine the absolute and relative quality achieved with the various codes (as affected by the variously modified parameters). Also, a test was run with a large number of observers to determine the relative recognizability of objects in one of the pictures when variously coded in color or monochromatically.

The results show that while the best monochromatic reproduction achievable in the experimental system requires a transmission rate of 5 bits per sample (with logarithmic quantization used), the best color reproduction in the same system (with the same luminance sample density) requires an average of 5.55 bits per sample. This is achieved by quantizing chrominance to approximately 1000 values and reducing the spatial density of chrominance samples to 1/18 of that of luminance. The results also indicate that the luminance sample density of a color picture can be reduced by a factor of from 1.5 to 18, or more, and still be equal in quality to the monochromatic reproduction, the amount depending on the subject matter and on the criterion used for comparison.

Two major conclusions were drawn from this study. (i) A normal monochromatic picture can be converted into a full color picture of the same apparent sharpness by transmitting additionally only a fraction of a bit per sample. (ii) For many purposes, inclusion of color may result in an over-all lower transmission rate requirement than would the same picture coded monochromatically; for some purposes, such as recognizing objects, this reduction can be substantial.

AUTHOR

TABLE OF CONTENTS

I. Introduction 1

II. Experimental Procedure 6

2.1 Basic System 6

2.2 Code Reduction 12

III. Results 17

3.1 Effects of Code Reductions on Appearance of Reproductions 17

3.2 Subjective Comparison Tests 24

3.3 Subjective Recognition Tests 28

IV. Conclusions 32

4.1 Discussion of Results 32

4.2 Summary of Conclusions 39

Acknowledgment 40

References 41

I. INTRODUCTION

Vision is the most comprehensive and versatile of human senses. Stated differently, vision, of all senses, has the largest capacity for gathering information for the human observer. Therefore, picture transmission, making wide use of the sense of vision, has, of all modes of conveying information to a human receiver, potentially the largest capacity for conveying such information.

The terms 'information' and 'capacity' are used here in their intuitive, nontechnical sense, although an analogy will be drawn later between them and the corresponding terms of statistical communication theory. Measurement of visual-information input in such statistical terms is extremely difficult and will not be attempted or referred to in this work. There will, however, be reference to 'information rate to the observer', to be explained as part of the mentioned analogy. The only quantifiable information rate that will be dealt with is the transmission rate through a physical channel of communication which is necessary to reproduce a picture as coded. This rate, which is of course measurable in the usual probabilistically defined units, is in practice also very high, compared with that necessary to stimulate other human senses. It is the minimization of this rate, by efficient coding of pictures, that is the main objective of the present study, as it has been with many previous studies.[1-6]

By 'coding of pictures' we mean selecting, from the original picture or scene, a finite number of quantities to represent it (or the information to be conveyed) and reconstructing from them, after transmission through the physical medium, an image to be presented to the observer. Coding of pictures as defined here should not be confused with the coding and decoding of signals for transmission over the physical medium, as treated by statistical communication theory. Inasmuch as a physical medium is a part of a picture transmission system, however, such statistical coding may be considered part of the problem of efficient picture coding. This aspect of picture coding is not treated in the present work; any calculations of transmission rates, that is, minimum physical channel capacity, are based on zero-order statistics.

The concept of picture coding can perhaps be illuminated further if we consider the whole chain, from the original picture through the eye to the consciousness of the observer, as analogous to a conventional transmission system. We then regard the whole physical system through the reconstructed image as an information source plus coder, the psychophysical relationships between this image and the sensors of the eye as the medium of transmission (or channel), the 'eye' (this term to be understood in the sequel to include some or all levels of the visual nervous system) as the decoder, and the observer's consciousness as the receiver (or information sink). Figure 1 shows this concept schematically. A block marked 'information source', customary in diagrams of a communication system, is here implicit as a variable portion of the original picture, representing those quantities in the picture selected by the 'coder' for transmission. (Our concept could be extended to include systems in which the original data are not

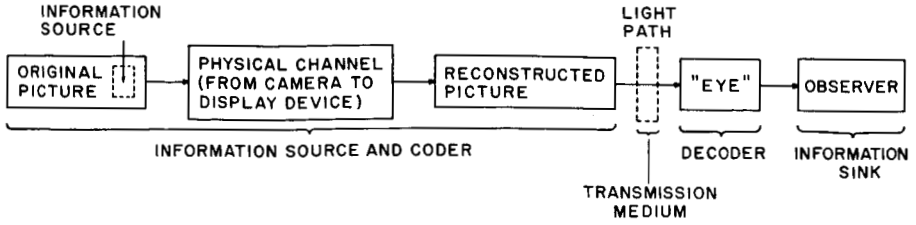


Fig. 1. Picture-transmission system as a communication channel.

visual in character but are displayed visually. The 'original picture' would then be replaced by the appropriate data source. Some of the conclusions of this report may still be valid for such systems.)

Devising a code for pictures consists of three major steps: (i) choosing what may be regarded as a measure of the amount of information conveyed to the observer, namely a criterion for the subjective quality of the reconstructed picture, (ii) deciding on the modality of the reconstructed picture, that is, which visual senses should be stimulated by the displayed image and in what ways, and (iii) selecting the minimum data to be extracted from the original picture that must be transmitted in order to reproduce the picture in the particular modality with a desired quality, as judged by the chosen criterion. Step (i) may involve one or several levels in the visual nervous system of which the observer could possibly be conscious, beginning with elementary sensory phenomena such as lines or shades and ending with total perceptions, such as depth, lighting or recognizable objects. One can specify a definite criterion at a certain level (for example, sharpness or recognizability of objects), or one can specify a more vague criterion relating to aesthetic or other personal values (for example, beauty or interest), thereby leaving it up to the individual to decide on the particular criterion or combination of criteria to be applied.

In line with our analogy, the various levels of cognition in the visual system could be conceived of as a series of abstractions, the flow from one level to the next being analogous to 'recoding' of information from one alphabet into another. In these terms, establishing a criterion of quality at a particular level amounts to selecting a certain class of 'words' in the corresponding code and asking whether or not they are present and, if so, whether or not they are uniquely decodable. To meet such a criterion, sufficient information must be supplied from the preceding level. In general, it may be assumed that such information is coded with some redundancy. Moreover, the preceding levels may transmit some information that does not contribute at all toward the recoding of the class of words of interest. All of this information is obtained, in turn, by recoding information from a preceding level, and so on down to the first sensory level.

Many relations between the displayed image and the lower sensory levels, known as psychophysical and corresponding in our model to the first few stages of decoding, have been determined experimentally. They give us basic indications of steps (ii) and (iii) for the correct way to code a picture, namely those (and only those) elements to be included

in the image which evoke significant responses in the eye. Thus, for example, sensory quality of acuity determines the maximum spatial frequency of brightness variation required in the image; any higher frequencies will not add to the sense of acuity and will only be wasteful of physical channel capacity that will obviously have to be provided to create them.

It is indeed on such psychophysical relations that current practical picture codes (such as the standards of commercial television[7,8]) are devised, the exact quantitative specifications having finally been determined by direct subjective testing. We shall refer to them as 'elementary codes'. Thus, an elementary code is designed for a hypothetical 'decoder' consisting of the first few recoding stages of the 'eye', namely, up to the level at which the elementary sensory criteria have been set. More efficient codes may be devised by expressing the elementary code in terms of different physical quantities, for example, spatial brightness differences, and by exploiting the statistical relationships among them.[9,10] Psychophysical relations between these, as well as other, new elements and the lower sensory levels indicate the possible further efficiencies[1-6] whereby physical picture elements may even be characterized by such sensory terms as edges, lines, and corners.

Even greater efficiencies are possible if the criteria are established at higher cognitive levels, ideally at those involving perception. This can be demonstrated, on the basis of earlier discussion, as follows: Because of the possible redundancy in the transmission of information to the 'critical' level(s), we should be able to eliminate or alter some words at some level without affecting the decoded information. In particular, we can eliminate or reduce certain groups of words (or alphabet symbols) in the lower sensory levels by appropriately modifying the displayed image. Thus, for example, we can eliminate all color information, fine detail information or temporal change information without affecting the usefulness of the picture for certain purposes. In general, there may be several alternative ways to perform this reduction, resulting in equivalent but different codes. One of these must be the elementary code. It is likely, however, that some of the other ones will be more efficient, since the elementary one was not designed to be so.

It is my thesis that the use of color as an element of the displayed image provides for one alternative code, which for many purposes (that is, for many criteria of quality) may be more efficient than any code using monochromatic display. Color is thus regarded as not necessarily constituting additional information, burdensome on the channel, but rather and primarily as an additional information carrier, that is, alphabet, adding to the potential over-all coding efficiency.

The basic psychophysical facts concerning color vision are incorporated in the science of colorimetry.[11,12] Its main postulate is that any color can be matched by the linear addition of three properly chosen colors, called primaries. Thus, three quantities are necessary and sufficient to specify the color of any spot of uniform surface (these quantities being the brightness of the three primaries or any reversible

transformation thereof). A monochromatic color requires only one such quantity (for example, brightness). On first glimpse it would therefore appear that the transmission of color pictures would require three times the rate of transmission of similar monochromatic pictures. This implies a naive code having three quantities for each one in the corresponding monochromatic picture.

A reduced code for color television, based on rough quantization of chrominance, has been proposed by Valensi.[13] The N. T. S. C. standards for color television[8] specify the same nominal bandwidth as monochromatic television, mainly by reducing the bandwidth of the chrominance, as well as by making use of statistical properties of the luminance signal. The codes reported on here incorporate the idea of chrominance quantization and extend the chrominance spatial bandwidth reduction to two dimensions.

It will be shown in this report that (a) the mere addition of color information to a typical monochromatic picture requires an increase in the transmission rate by only a small fraction to maintain at least the same quality, as judged by any criterion, and, moreover, (b) for some criteria, the addition of color actually can result in a lower transmission rate for the same quality.

Result (a) is in some sense an extension of the elementary code: It too is based on some fundamental psychophysical relations, namely the relatively low acuity to spatial frequencies of color variations and the limited chrominance discrimination ability, and it too was determined by direct viewing of typical pictures with criteria placed at the lower levels of the sensory chain. Specifically, two separate coding parameters relating to color were investigated: One was the highest two-dimensional spatial frequency of the chrominance, and the other, the number of different chrominance values, that is, 'colors'. They correspond to the similar two parameters considered in the elementary monochromatic code. The criterion of quality was simply the lack of any noticeable distortions above those present in the naively coded color picture.

Result (b) is suggested by result (a): If it 'costs' so little extra transmission capacity to add all of the perceivable color information, would it not be possible to obviate this extra capacity by slightly reducing some other information while keeping the total information conveyed to the viewer in some sense increased over what he obtains from the monochromatic picture? An alternative way of asking this question, and one more amenable to testing (in the absence of quantitative measure of observer's information input), is: If we assume that the addition of color to a picture increases its informational value to the observer, that is, enhances the picture quality, how much can we reduce other elements of the picture before it is degraded to its original quality, and would this not lead possibly to a lowered over-all required capacity? The elements that were reduced in my tests were the two fundamental parameters of monochromatic pictures, namely the two-dimensional spatial frequency and the number of brightness levels. The criteria applied involved higher levels of the visual chain, and will be explained in the description of the experimental procedure.

My study is concerned only with the case in which the color of the displayed image

is as close as possible to that of the original picture – which is usual and most desirable for many purposes. Other cases may involve deliberate enhancement or shift of color or original scenes consisting of infrared or x-ray images or nonpictorial data. The displayed image therefore will also be called 'reproduction'.

II. EXPERIMENTAL PROCEDURE

2.1 BASIC SYSTEM

The experiments consisted essentially of creating variously coded reproductions of two original pictures on color-film transparencies and projecting them for viewing by observers. The two pictures were of typical natural scenes, and their coded versions are reproduced in Fig. 4 ("fruit") and Fig. 5 ("girl").

The coding was carried out by means of the digital picture recorder and reproducer, specially built for this and similar research by the Picture Processing Group of the Research Laboratory of Electronics, M.I.T., and the IBM 7094 data processing equipment at the Computation Center, M.I.T. A full description of the recorder-reproducer is available.[14] The essentials of its operation, however, will be described here as part of the coding procedure.

All of the codes were obtained by reducing, in various ways, one fundamental code, which therefore presents an upper bound on the quality of pictures 'transmitted' over my system, and, on the other hand, an upper bound on the equivalent rate of transmission. The reductions were carried out by appropriate programming of the IBM 7094 data processing equipment and, in effect, recoding in terms of the fundamental code (the various codes being thus only simulated on my system). The fundamental code was established primarily by appropriate settings of the recorder-reproducer.

The basic procedure is shown schematically in Fig. 2. The source picture, an Ektachrome transparency, is sampled in an array of 256 × 256 sample points, by scanning it with a discretely positioned (digitally controlled) flying spot, imaged from the face of a cathode-ray tube. The light transmitted by the transparency at each sample point is split into three paths, each path passing through a Wratten filter and onto the cathode of a multiplying phototube. Each of the three filters – red, green, and blue – has been chosen so as to have a band sufficiently narrow to allow passage of the light modulated by only one of the corresponding dye layers in the transparency, and still pass enough light to induce a usable signal in the phototube. The currents obtained from the phototubes are amplified, the factors of amplification being so adjusted that the resultant three signals are equal when a white or neutral gray area is being scanned. They are thus proportional to the intensities of the three hypothetical primaries that would combine to create the image of the slide as sampled, and will be designated in the sequel as the R, G, and B signals, respectively. These signals are gated sequentially, in the order B-G-R, amplified, and passed through a logarithmic attenuator. The output of the attenuator is proportional to the logarithm of its input signal and, therefore, inversely to the density of the dyes in the slide. This facilitates correction, at a later stage, for unwanted absorptions in the dyes (that is, cross modulation of color signals), which are proportional to their densities. The signal is then integrated over the duration of the sampling pulse and quantized to 1024 levels by an analog-to-digital converter. The

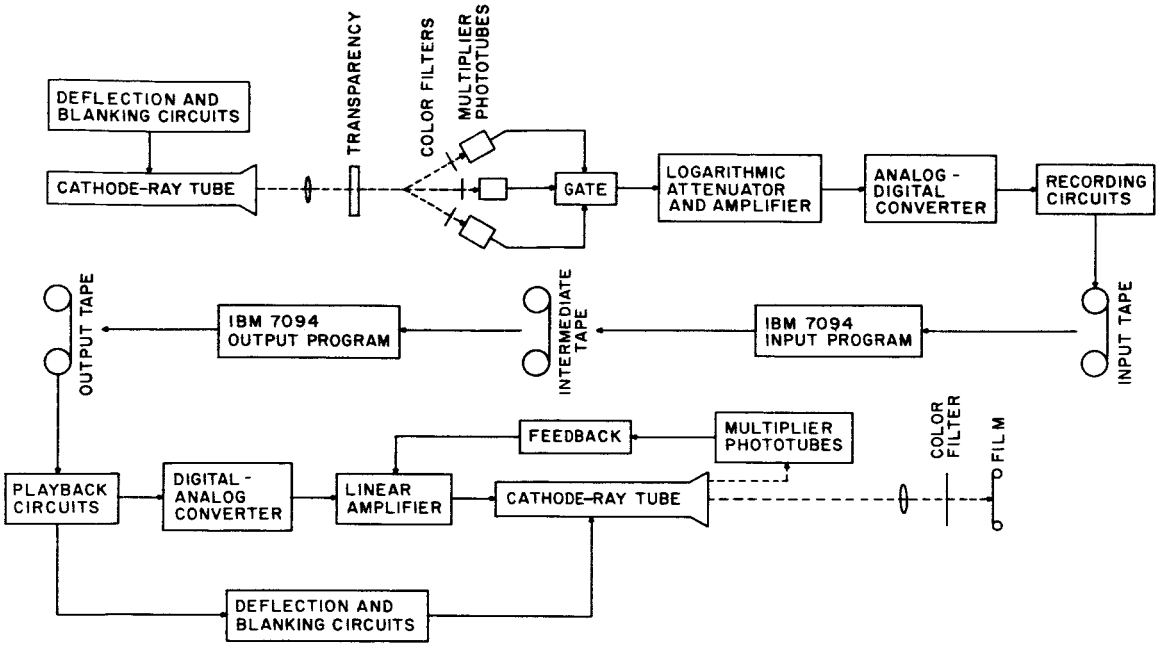


Fig. 2. Block diagram of the experimental arrangement.

digital output of the latter is recorded on magnetic tape, which becomes the input tape to the IBM 7094 computer.

A program called the 'input routine' performs the following operations in the IBM 7094 computer: Data (representing sequentially the logarithm of the intensities of the three primaries of successive samples) is read from the input tape. Corrections for dye absorption are made for each of the three signals by using formulas of the form

$$G' = (1+m_b+m_r)\,G - m_b B - m_r R,$$

where G, B, and R are the recorded logarithmic signals, G', is the corrected logarithmic green signal, and m_b and m_r are the ratios between the attenuation by the magenta dye of blue and red light, respectively, and that of the green light. Similar ratios are defined for the cyan and yellow dyes. These ratios have been determined by direct measurement and found to be approximately

$$c_g = 0.14 \qquad y_g = 0.09$$

All other ratios are near or equal zero.

This operation corresponds to 'masking' in photographic color reproduction and renders the colors in the resultant image truer to the original scene.

The signals are then converted back to the linear intensity scale (by using the inverse characteristic of the logarithmic attenuator), after which the picture information is recoded in terms of luminance and chrominance quantities. The luminance is the sum of the intensities of the three primaries after being weighted by the standard

ICI luminosity factors and is regarded as being (or carrying all the information of) the monochromatic component (or version) of our pictures. [In monochromatic reproduction various chromatic weighting curves may be used (for example, those of orthochromatic or panchromatic films with various filters), representing, in our terms, different codes (requiring the same transmission rate), some of which may produce superior results over those of others. Indeed, some such monochromatic pictures produced by our system were preferred by some observers to those rendered by luminance. The differences, however, were, in general, inconsequential for our purpose. From a practical point of view, monochromatic reproduction of luminance is simpler on our system.] The chrominance comprises all information not included in luminance and is expressed in terms of two independent variables. These variables are represented by the X and Y coordinates of a plane having a "Uniform Chromaticity Scale" (UCS) as proposed by Breckenridge and Schaub[15] (see Fig. 3). Chromaticities on this plane are obtained by

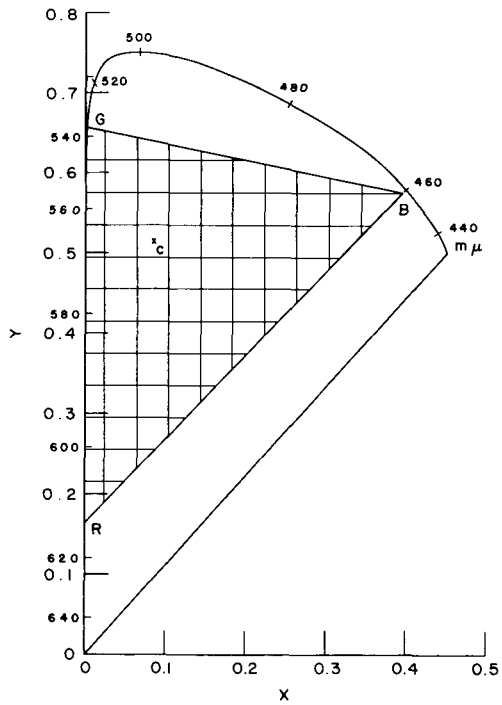


Fig. 3. Chromaticity quantization of step size . 04 and triangle of reproducible chromaticities on a Uniform Chromaticity Scale.

oblique projection from that of the standard ICI chromaticity diagram, but the UCS has the property that chromaticities having equally noticeable differences are represented by approximately equidistant points. The use of this scale is convenient for efficient quantization of chrominance.

All possible chromaticities transmitted in my system fall within a triangle whose

corners represent the chromaticities of the three hypothetical primaries. These can be calculated from the spectral characteristics of the dyes (as corrected), the phosphor of the scanning cathode-ray tube, the filters, and the cathodes of the phototubes; if complete separation of dye layers into signal is assumed, the dyes alone determine the chromaticities. Since the exact absolute chromaticities in the original scene were not of great importance for the purpose of our study, however, and in order to ensure the retransformability of all chromaticities into positive primaries, the chromaticities of the hypothetical recording primaries were conveniently assumed to be identical to those of the reproduction primaries. These are marked in Fig. 3 and their calculation will be explained later. The transformation of the recording primaries B, G, R into luminance, L, and chromaticity coordinates, X and Y, follows the formulas

$$L = .0668\,B + G + .4276\,R$$

$$X = \frac{1}{D}(.15834\,B + .00416\,G + .00002\,R)$$

$$Y = \frac{1}{D}(.23103\,B + .65691\,G + .07847\,R)$$

$$D = .40275\,B + G + .47942\,R$$

(For computer efficiency, all multiplications were carried out by adding logarithms before converting them as described.)

All operations up to this point are identical for all codes investigated. Therefore, they are carried out only once for each source picture. The quantities L, X, and Y are recorded on an intermediate tape and used subsequently as the input data for all variable-coding operations.

The sequence of computer operations from reading the intermediate tape to writing the output tape comprises the 'output routine'. It starts with one or several of the variable-coding operations (which, as mentioned, represent certain reductions from the fundamental code), all performed on the X, Y, and L quantities, as described in detail in section 2.2. To obtain a picture by the fundamental code, these quantities are left unchanged.

Next, X, Y, and L are retransformed into primary quantities B, G, and R. The latter are the signals for the intensities of the virtual primary color sources that combine additively to produce the final image (as will be explained). The retransformation formulas are

$$B = \frac{L}{d}(-.23646\,X + .00197\,Y - .0003133)$$

$$G = \frac{L}{d}(.07915\,X - .07590\,Y + .0124203)$$

$$R = \frac{L}{d}(.03353\,X + .15667\,Y - .1030540)$$

$$d = .077690\,X - .008782\,Y - .0316665$$

These are exactly the inverse of the transformation formulas given previously, by virtue of the assumed identity of recording and reproduction primaries. The quantities B, G, and R are then written on the output tape, in the same format as on the input tape. Note that they now represent light intensities linearly (not logarithmically).

The reproduction of pictures proceeds as follows: The output tape (as recorded by the computer) is played through the tape unit of our recorder-reproducer. The digital data on it are converted into an analog signal, which is amplified and applied to the scanner cathode-ray tube circuit, with the purpose of modulating the spot intensity. Since we effectively reproduce colors by the addition of primaries, we wish the spot intensity to be as linearly related to the signal as possible. For this purpose a servomechanism system has been built around the cathode-ray tube, in which the light from the spot on its face is monitored by a multiplier phototube. The signal from the phototube is compared with the input signal (which comes from the tape), and the difference between them is highly amplified, integrated, and fed to the grid of the cathode-ray tube. A maximum deviation from linearity of approximately 10 per cent was measured in the useful range of the system.

The spot is deflected in exactly the same digitally controlled manner as during recording, synchronized by the data read from the tape. A 'wobble' in the form of a Lissajou pattern, of peak-to-peak amplitude approximately one-half the distance between sample points, is added to the deflection. Also, the spot is defocused electronically until neighboring points appear to the eye to begin to merge. The wobble and defocusing are the equivalent of an aperture function with which sampled data are normally convolved to eliminate the high-frequency bands introduced by the sampling process. The resultant spatial spreading of the sample pulses also helps to keep the light intensity within the useful range of the film (otherwise, sharp exposed-to-saturation dots will appear on the film, surrounded by dark areas). The spot is displayed at the intensity level specified by the signal for a certain duration (the same as during recording) and blanked off between samples. Actually, because of the integrator in the servomechanism circuit, the total light energy over a display pulse, rather than the intensity, is made proportional to the signal, since this is really the quantity that determines the stimulus to the eye or to the photographic film. In time, the spot-by-spot display thus creates on the face of the cathode-ray tube an image of the desired reproduced picture, which, in turn, is projected by a camera onto a film for permanent record and subsequent viewing. The three primary color images are played back separately and are simply superimposed on the film. Since the three signals B, G, and R are recorded on the tape sequentially for each point, the spot is kept blanked off for all but one out of every group of three samples -- namely for the one corresponding to the primary being photographed.

The colors are physically 'created' by placing Wratten filters over the lens (No. 47B for blue, No. 74 for green, and No. 29 for red) and exposing high-speed Ektachrome film. The chromaticities of the three effective primaries have been calculated from the

specified spectral characteristics of the P24 phosphor of the cathode-ray tube and of the filters. No consideration has been given in this calculation to the characteristics of the film; the subtractive nature of its colors and its nonlinearities would have rendered the chromaticities of real colors difficult to calculate and, moreover, not constant over the brightness range; also, at one time the use of several different films was considered. Thus, the superimposed images on the face of the cathode-ray tube, as seen through the filters, is considered to be the displayed image of the received (that is, coded) picture, whereas the film merely carries a photographic record of it, in the same manner and with the same fidelity as of any natural scene.

For correct exposure of the film, three quantities were determined experimentally: the relative lens opening for the three colors, the absolute lens opening, and the light level of the spot for zero signal (black). The relative lens openings were found to be $\frac{1}{2}$ stop from one another for neutral gray rendition of equal intensities of primaries. The absolute lens opening was determined after establishing the exact maximum spot intensity (the maximum that the cathode-ray tube servomechanism circuit could handle at a particular focus setting and 25 kv ultor voltage) and its rise-time characteristic (as controlled by a potentiometer in the servomechanism circuit). Optimum pictures were found to be produced at these settings with lens openings of f 4, f 4.5, and f 5.6 for the "girl" and at f 4.5, f 5.6, and f 6.3 for the "fruit" scene. The black level (which effectively determines the available contrast of the reproduced picture and is controlled by another potentiometer in the servomechanism circuit) was found to be optimal at the extinction point of the spot (as it should be) for the "fruit" and slightly above it for the "girl" (because of the abundance of dark detail that would have fallen within the 'toe' portion of the film characteristic).

The size of the image on the film is 36 mm square, and the developed film is cut into transparencies and mounted in standard size cardboard frames for direct viewing or projection. In the subjective tests the slides were usually projected to a size of approximately 20 in. square and viewed from a distance of approximately 8 feet. At such viewing the spots are on the point of seeming to merge, as they subtend a visual angle of approximately 3 minutes.

Every precaution was taken to keep the whole process of picture reproduction, from digital signal to the final viewing, as uniform from one picture to another as possible, so that the only noticeable differences would be due to the programmed code reductions. Unfortunately, some unwanted variation did occur between reproductions, notably in exact color balance (resulting from small differences in relative primary exposures) and contrast (because of changing black level resulting from thermal drift in the circuits). Most of the slides were developed by a commercial laboratory in a single batch. They showed such a marked shift in color balance from previous batches (having more normal balance), that a compensating filter (CC40B) had to be used in their projection (and in the reproductions in this report). The unwanted variation between pictures may have had some effect on the judgment of the viewers, but this effect was significant only where

the noticeable difference caused by coding was very small.

There is, of course, no point in comparing the reproduced pictures with the original transparencies, because of the severe degradation introduced by even the fundamental coding. The main factor here was the very limited number of samples, which reduced fine detail and edge sharpness. Another source of degradation, which was not inherent in the code but was due rather to wrong settings of parameters in the system, was a considerably higher contrast in the reproduction. Apparently the assumed intensity range of 1:32 for the original (which was taken to be the basis of the logarithms during the computer conversion of the signals) was too high. This resulted in dark colors being too dark, while colors in bright areas tended to become "washed out" (highlights even becoming completely white) because of saturation of the film. This degradation was more harmful to the color reproductions than to the monochromatic ones. Therefore, any conclusions drawn in favor of color could have been even stronger, and my conclusions would have been qualitatively unchanged. A series of black dots in the reproduction, caused by burned spots on the face of the cathode-ray tube, did not seem too disturbing to the viewers. The color prints used for illustration in this report were reproduced photographically from the transparencies.

2.2 CODE REDUCTION

In testing the first premise of this report, concerning the coding of the chrominance information, two quantities, the spatial density of the chrominance samples and the number of quantum steps of chrominance, were reduced to various degrees, separately and in combinations. Similarly, in dealing with the second premise, concerning the reduction of over-all information, the sample density and number of quantum steps of the luminance were reduced, to various degrees, separately, in combination by themselves, and combined with the reductions in chrominance.

All such reductions were carried out in the computer as part of the output program. All reductions involving quantization were written as interchangeable subroutines, named UNPAK (since they included unpacking of data from word format on the intermediate tape). All reductions involving number of samples were programmed as interchangeable subroutines, named PROCS (for 'process'). Thus, various combinations of these quantities were achieved by simply assembling various UNPAK and PROCS subroutines with the output routines.

Reduction in the number of samples of a picture quantity (for example, luminance or chrominance) that must be transmitted is always possible, according to the sampling theorem, if the highest spatial frequency in the original picture is first reduced proportionally, with a corresponding reduction resulting in the reproduced picture. Reduction of the highest spatial frequency is called lowpass spatial filtering. We are dealing here with the two-dimensional case, an extension of one-dimensional spatial filtering, which is in many respects analogous to filtering of temporal functions. (The subject of two-dimensional sampling and filtering of pictures has been investigated experimentally

by Tretiak.[16]) This filtering is conveniently done in the space (rather than frequency) domain by convolving the picture with an 'aperture function', which is the two-dimensional Fourier transform of the filter function. The ideal lowpass-filter function would be constant within a circle that represents the cutoff frequency and zero outside it. The corresponding aperture is a radial Bessel function. To save computation time, however, the actual aperture function used in simulating code reduction was simply constant within a circle and zero outside it. This represents a mere approximation to lowpass filtering, but it can be justified as adequate for our investigation by pointing out that the high frequencies that are thus allowed to pass are comparatively small in magnitude and, moreover, are shifted in phase and therefore introduce distortions that tend to push our estimates of acceptable code to the conservative side.

The filtering program consisted then of replacing each sample by an average of all of the samples within a circle around it. The number of samples within this circle represents the factor by which the number of samples that must actually be transmitted can be reduced without affecting the quality of the received picture (within the approximation discussed above). Five degrees of reduction were made possible, the number of samples within the aperture being 5, 9, 13, 21, or 37. Because of some inherent lowpassing in the recording and reproducing systems, the unfiltered picture was judged empirically (by various subjective estimates) to be equivalent to one filtered by an aperture of the type discussed above with a size containing two samples. Thus, the following reduction ratios in the sample number were tested: 2:5:9:13:21:37. A particular filtering operation on a picture is designated as to luminance (L) or chrominance (XY) and the aperture size. For example, a picture whose chrominance is filtered with an aperture of a size containing 21 samples but with no filtering of the luminance is designated as XY21 (with L2 being understood); a picture in which both luminance and chrominance are filtered with an aperture of size 9 is designated as L9 (with XY9 being understood, since there is obviously no point in ever filtering the chrominance less than the luminance).

Quantization of luminance was carried out on a logarithmic scale to ensure equal noticeability of steps over the entire luminance range. The roughest quantization divided the range into 8 such logarithmically related levels (which was very easily programmed by counting the number of significant places in the binary representation of the luminance). Finer quantization was achieved by allowing 1, 3 or 7 intermediate levels (linearly spaced) between these 8 basic levels, to obtain totals of 14, 27, and 54 levels, respectively, over the luminance range. Actually, a printout of the distribution of the samples over the luminance range for both pictures indicates that they were confined almost entirely to approximately two-thirds of the range (possibly because of some flaw in the recording process). Therefore, it would be more realistic to assume that the reproduced luminance range was divided into only 5, 9, 18, and 36 levels, respectively. The corresponding basic rates of transmission would be approximately 2, 3, 4, and 5 bits per luminance sample.

Chrominance was quantized very simply by dividing each of the X and Y chromaticity coordinates into equal steps, thereby dividing the chromaticity plane into equal squares. Each square was represented by the chromaticity of its center. The raster of boundary lines was positioned so that the white (Illuminant C) point was in the center of a square and therefore all near whites appeared to be white. Such division of the uniform chromaticity scale results, presumably, in approximately equally noticeable steps, but this was not tested directly. Aside from this fact and the placement of the white point in the center of a square, this division was entirely arbitrary. Division of the chromaticity plane by lines of constant hue and lines of constant purity would be in some sense more 'natural' and would probably result in better pictures (for equal over-all number of quantum areas) because surfaces in natural pictures tend to have constant hue but gradually changing purity and intensity. I judged, however, that the added quality thus achievable is not worth the considerably greater complication in programming and additional computer time required. (In most of my chromaticity-quantized pictures, artificial contours of purity steps, rather than hue steps, were the most disturbing effects, just as would have been the case with constant-hue lines quantization.)

The following quantum step sizes, as measured on the UCS, were employed, corresponding to increasing roughness of chromaticity quantization: .01, .02, .04, .056, .08, and .10. The triangle of reproducible chromaticities was thus divided into approximately 1040, 270, 70, 32, 22, and 16 squares, respectively. A typical case, for step size .04, is depicted in Fig. 3. The fractional squares along the sides of the triangle are, strictly speaking, complete quantum areas by themselves, so that the total number of areas should be larger than noted above. For most practical pictures (and certainly for ours), however, the colors along these sides seldom appear, so that in any practical code with such a rectangular raster used for quantization the smaller of the fractional squares could be merged with their neighbors. Furthermore, a diagrammatic printout of the distribution of the chrominance samples in both of our pictures showed heavy concentration around the white (or gray) point and along a few thin radially directed regions. (This distribution suggests the possibility of using statistical coding techniques in practical applications for greater efficiency.)

If we add luminance to the chromaticity plane as a third dimension, we obtain a coordinate system for the whole color space. All reproducible colors are enclosed within a surface in this space having the shape of a tent, with a triangular base, supported by seven poles located at the chromaticity coordinates of the three primaries, the three points representing equal mixtures of two primaries, and white. The lengths of the poles are proportional to their luminosities, that is (in arbitrary units) 1 for green, 0.428 for red, 0.067 for blue, 1.428 for red-green (yellow), 1.067 for green-blue, 0.495 for blue-red, and 1.495 (=1 + 0.428 + 0.067) for white.

By quantizing both chromaticity and luminance, the space is divided into square-shaped 'bricks'. If each 'brick' is assigned a code word, greater efficiency results than when chromaticity and luminance are coded separately (because not all possible

combinations are within the reproducible space). Additional efficiency may result from such combined coding if statistical distribution over the space is brought into consideration. At the beginning of the investigation it seemed possible that greater efficiency could also result from consideration of the psychophysical relations between luminance and chromaticity. In particular, I thought of taking advantage of the fact that the lower the luminance the less discernible the chromaticity, so that coarser quantization of the latter may be indicated then. Indeed, I made several pictures with both chrominance and luminance quantized, in which the fineness of the chrominance quantization increased for each higher level of luminance. Three degrees of luminance-dependent chrominance quantization, called 'fine', 'coarse', and 'rough', were employed. Their construction is shown in Table 1.

Table 1. Luminance-dependent chrominance quantization.

Chromaticity quantization step size	Luminance level at which quantization applies		
	fine	coarse	rough
.01	8	—	—
.02	7	8	—
.04	6	7	8
.056	5	6	7
.08	4	5	6
.1	3	4	5
monochromatic	1-2	1-3	1-4

The approximate number of 'bricks' within the reproducible space of colors was calculated for the 'fine', 'coarse', and 'rough' luminance-dependent chromaticity quantizations to be 93, 29, and 12 times the number of luminance levels, respectively. The approximate total number of quantum 'bricks' within the reproducible space was calculated similarly for independent chromaticity quantization of step sizes .01, .02, .04, .056, .08, and .1 to be 520, 135, 35, 16, 11, and 8 times the number of luminance levels, respectively. The total number of quantum 'bricks' (or values) for several such combined quantization schemes is given in Table 3.

Coding based on the combination of quantization of chrominance and luminance, as discussed above, reveals itself to be of no great value for the following reason: As our results indicate, a much more substantial reduction in transmission rate can be achieved by greatly reducing the density of chrominance samples. Thus, the majority of samples would be of luminance alone, and the 1 or 2 bits that can be saved on the few combined samples would amount to very little when averaged over all samples. One could perhaps

do somewhat better by making the luminance quantization dependent on the proximate (or interpolated) chrominance and coding accordingly, bringing into consideration again the boundaries of reproducible colors. I have not tried such a scheme, however.

III. RESULTS

As discussed in Section II, the four basic parameters governing the coding of our color pictures (namely, the quantization and the density of samples of the luminance and of the chrominance), were reduced to various degrees from their fundamental values, singly and in combinations, by simulation with appropriate computer programs. Each such value or combination of values was considered a code and was usually applied, in one computer run, to both test pictures, and required an average of one minute of computer time per picture. The output tapes of several such runs were then played as described in Section II to reproduce the coded pictures on Ektachrome film (size 120). It took an average of 7 minutes to expose each frame (including tape rewinding and filter change). The commercially developed and mounted transparencies were then viewed either directly on a light box or by projection.

Approximately 150 transparencies were made during the earlier phases of the work for two purposes: first, to experiment with and adjust the various settings and parameters of the systems, particularly those of the reproducing circuits and optics, and, second, to obtain some preliminary ideas about the effects of the various codes so as to ensure being in the correct range.

Approximately 100 transparencies were then made, as a main 'production run' for serious evaluation. Eighteen of these are reproduced here in the color plates (Figs. 4 and 5) for illustration. Except for the effects of the codes themselves, all of these transparencies were fairly uniform; minor variations that did appear and their causes were discussed in section 2. 1. Also, several obviously bad frames of film were made because of tape errors, externally and internally produced electrical disturbances (all of which usually resulted in loss of synchronization of the scan), and erroneous electrical and optical settings.

The coded pictures were evaluated in three phases: (1) my own critical observation of the effects of the codes on the appearance of the reproductions, aided by the observations and comments of other individuals (including some made during the tests mentioned below); (2) evaluation by 4-6 subjects of the relative merits of various pictures within certain groups, compared in twos; (3) testing of groups of subjects on the recognizability of objects in various pictures. Each of these phases and the results it yielded will now be discussed.

3. 1 EFFECTS OF CODE REDUCTIONS ON APPEARANCE OF REPRODUCTIONS

The apparent effects of the reductions in the various parameters, first singly, then in combinations, are discussed.

a. Lowpass Filtering of Chrominance

Upon direct viewing of transparency or normal-distance viewing of projection, no difference is apparent between the fundamental picture (reproductions 1, 10) and XY21

and XY37 (reproductions 2, 11), except a slight blurring of the geranium flowers. An equal filtering of luminance is, by way of demonstration, markedly noticeable (reproduction 7). On close examination of the projected image, some blurring of chrominance was noticed across edges at which the luminance contrast was relatively small. It was interesting to observe, in this connection, that because of the roundness of most objects in the picture, there was usually a dark margin along their edges, which either increased the contrast with the adjacent surface or merged with the latter's dark margin; in either case, this probably contributed greatly to the unnoticeability of the chrominance blurring. The effect of chrominance blurring was particularly evident (again on close examination) where a narrow stripe of one color was flanked by another color, in which case the former changed its apparent chrominance halfway toward the latter's (for example, if they were complementary it became gray).

b. Chromaticity Quantization

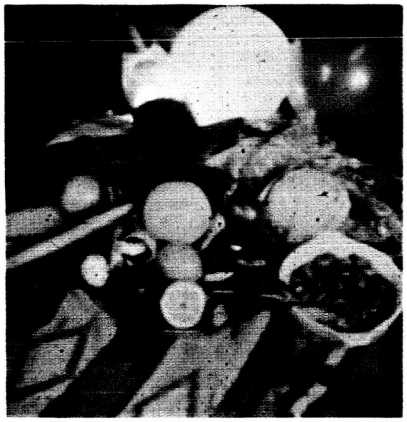
Pictures with chromaticity quantum size .01 are indistinguishable within the accuracy of the system from the fundamentally coded ones. With quantum size .02 (reproduction 5), some artificial contouring begins to appear in large surfaces, such as the balloons, grapefruit, and face; color shift is negligible. Quantum size .04 causes more severe contouring and also some color distortion (whereby, for example, light green in the balloon shifts to cyan, and the light yellow in the blanket shifts to white). At quantum size .056 (reproductions 6, 14) the color distortion becomes more prevalent (for example, orange shifts to yellow). At quantum sizes .08 (reproduction 15) and .1 the contours and color distortions become severe, so as to even 'break up' shapes of objects. Curiously enough, with the values .08 and .1, the quantum squares evidently happen to be so placed in the chromaticity plane that the entire face of the girl falls within one square and therefore is reproduced with a single chrominance value, which eliminates some of the objectionable contouring appearing in finer quantizations (but also renders the face less 'lifelike'). Likewise, at .08, the entire hair is reproduced in a single chrominance value.

c. Lowpass Filtering of Luminance

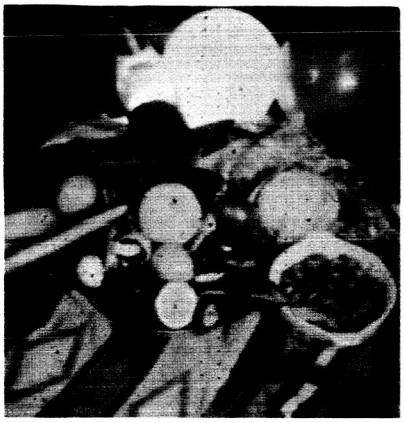
Increased blurring is evident as the aperture size increases (compare reproductions 1, 4, and 7 or 10 and 13), at a subjectively commensurate rate. The difference between L5 and the fundamental, however, warrants an assumed aperture ratio of 3:1 or 2:1, rather than 5:1. Consequently, the fundamentally coded reproduction is estimated to be the equivalent of L2.

d. Luminance Quantization

Only quantizations to 18, 9, and 5 levels were carried out singly. Quantization of 36 levels was carried out in combination with dependent chromaticity quantization. From the latter's reproductions (for example, reproductions 9 and 18), however, one can glean



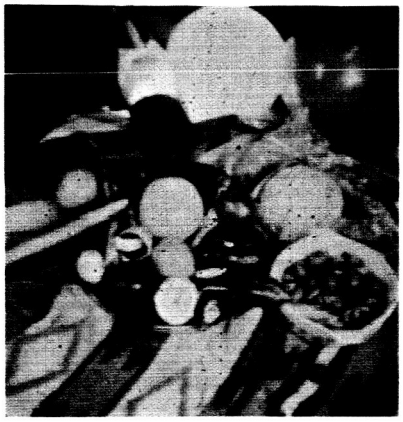
Reproduction 1
Fundamental (L2)



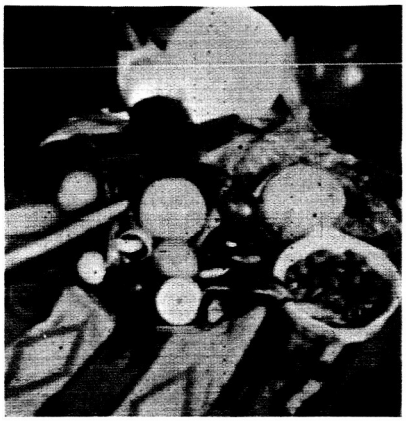
Reproduction 2
Chrominance low-passed XY37

Reproduction 3
Monochromatic (luminance)

Reproduction 4
Luminance low-passed L9



Reproduction 5
Chromaticity quantized . 02



Reproduction 6
Chromaticity quantized . 056

Reproduction 7
Luminance low-passed L37

Reproduction 8
Luminance quantized 18

Reproduction 9
Luminance quantized 36 and chromaticity quantized 'fine', XY21

Fig. 4. Reproductions of "fruit" picture variously coded.

Reproduction 10
Fundamental (L2)

Reproduction 11
Chrominance low-passed XY37

Reproduction 12
Monochromatic (luminance)

Reproduction 13
Luminance low-passed L13

Reproduction 14
Chromaticity quantized . 056

Reproduction 15
Chromaticity quantized . 08, XY37

Reproduction 16
Luminance quantized 18

Reproduction 17
Luminance quantized 9 and chromaticity quantized . 02

Reproduction 18
Luminance quantized 36 and chromaticity quantized 'coarse', XY37

Fig. 5. Reproductions of "girl" picture variously coded.

that no noticeable effects are introduced by luminance quantization to 36 (logarithmically related) levels (which concurs with well-known results). Quantization to 18 levels results in quite noticeable contours in large, as well as some medium-sized, smooth areas, and 9-level quantization causes considerable contouring and even some distortion in shapes. It is interesting to note that where severe quantization (even with 5 levels) distorts shapes of objects by eradicating parts of edge lines and adding erroneous ones, the presence of color tends to preserve the shape of the objects and to help distinguish one object from another, or avoid the perception of an object 'breaking up' into two.

e. Combination of Chrominance Quantization and Filtering

This combination was attempted with the anticipation that chrominance-quantized pictures will actually improve by lowpass-filtering the chrominance because the artificial contours would be blurred. As can be expected from our discussion of lowpass filtering of chrominance, however, this blurring is not very noticeable, except on close examination. It should be noted at this point that there is no guarantee that the luminosity and chromaticity of the hypothetical primaries of the transparency are the same as those assumed in the computer program; hence computed chrominance variation at constant luminance may well result in some variation in luminance in the transparency. This possibility explains some secondary observations, namely that chrominance contours do seem slightly sharper when not spatially filtered, that there is a slight 'crispening' of chrominance-quantized pictures over unquantized, and that this crispening is lost when chrominance is spatially filtered.

f. Luminance Quantization and Filtering

Luminance quantization and filtering was carried out with the same rationale as chrominance quantization and filtering (e). Artificial contours indeed become blurred to the same extent as do natural edges. The contours of 18-level quantization almost vanish with L21 filtering, but even without any filtering they are noticeable only in the balloons (especially the green one). With other quantizations, contours are noticeable, regardless of filtering.

g. Independent Chromaticity and Luminance Quantization

Several pictures were coded by quantizing both chromaticity and luminance in various combinations of quantum sizes. The noticeable effects are exactly the same as those reported for chromaticity quantization (b) and for luminance quantization (d). The contours caused by the two quantizations do not coincide, of course, but tend to prevail in the same areas and to be somewhat parallel (since gradations in luminance are usually accompanied by gradations in color purity, mainly as a result of the effects of specular reflection in the original scene).

h. Dependent Chromaticity and Luminance Quantization

A series of transparencies was made in which the chromaticity was quantized, with the quantum size dependent on the luminance of the point and, simultaneously, the luminance itself quantized, as explained in Section II. There were several combinations of 'fine', 'coarse', and 'rough' chromaticity quantization ranges with 36-, 18-, and 9-level luminance quantization. In general, the same observations apply here as did for independent chromaticity and luminance quantization (g), except that, as one would expect, there is a more complicated pattern of chrominance contours and shifts (again, noticeable mainly in large areas). Also, some of the chrominance contours coincide with some luminance contours (namely those corresponding to luminance steps at which the chromaticity quantum size is changed, which usually results in a change in the representative value of the chrominance), thereby rendering them more noticeable.

i. Dependent Quantization and Filtering of Both Chromaticity and Luminance

Several pictures were made in which all four parameters of the codes were modified from their fundamental values. Luminance and, dependently, chrominance were quantized in various combinations of quantum sizes; luminance was then lowpass-filtered with small apertures, while the chrominance was filtered with equal or larger apertures. The same observations about the effects of the filtering on the contours for quantization and filtering of chromaticity (e) and that of luminance (f) are valid here.

3.2 SUBJECTIVE COMPARISON TESTS

The real test for the quality of any picture-transmission system lies with the ultimate receiver, namely the observer, as discussed at length in the introduction. In order to determine the relative quality of the various codes simulated by our experimental system, I showed the resultant transparencies to several individuals and asked them to compare them with one another and state their preferences. The transparencies were projected on a matte white screen two at a time, side by side, by two identical projectors for as long as desired. The observers were tested alone, each in several sessions lasting 15-60 minutes. There were 6 observers altogether, although only 4 of them were shown all of the pictures. Two of them had been working on picture coding (monochromatic) themselves; one is an amateur photographer, and he and another observer are also interested in picture-transmission problems. The two other observers were females, with no relevant technical training. The comparisons were usually between pictures coded with one or several parameters modified to various degrees and those having other parameters modified.

Unless otherwise noted, the criterion by which the observers were asked to judge the quality of the displayed reproductions was, in the case of the "fruit" picture, the 'recognizability' of objects. For the "girl" picture it was variably stated as 'acceptability', 'freedom from annoying distortions' or 'which would you carry as your girlfriend's portrait?' Obviously, the criteria used for the "girl" involved aesthetic values

and gave the observers much more freedom. The variability between their judgments, however, was not enormously greater than that regarding the "fruit." Of course, even with the recognizability criterion, I had to rely on the observer's understanding of it and have no guarantee that other criteria did not in some way influence his judgment. Still, the agreement among the observers was generally good enough to make the results significant. In order to make some of these results more conclusive, a recognition test was devised and is described in the next section.

It is convenient to have one series of coded pictures as a 'standard scale' with which to measure, by comparison, all others. The series with the lowpass filtering of luminance was found to be most appropriate for this purpose because it spans the widest range of equivalent transmission rates with fairly evenly spaced values and because it has the most obviously monotonic variation of subjective quality. Thus any picture could be judged to have a quality equivalent to, say, L9 or L13 or, if intermediate, to L11. The best quality then is 'L2' (which is the fundamental code), and the worst, 'worse than L37'.

a. Monochromatic Transparencies – Spontaneous Judgment

Partially for its own interest and partially to obtain indications of bias of individual observers, the first test (before the observers became familiar with the pictures) was to compare a monochromatic (luminance only) reproduction of L2 and L5 (girl only) and of L9 with the standard series, with no criteria whatsoever being stated.

For "fruit" the results were: One observer (a monochromatic picture coder) found the monochromatic pictures to be equivalent to color of same L number; another observer (female), on the other hand, judged monochromatic L2 to be worse than color L37; the others placed monochromatic L2 between L9 and L21, and monochromatic L9 between L13 and L37. The equivalent on the average was

Monochromatic L2 = color L13

Monochromatic L9 = color L21.

For "girl" the results were less dispersed.

Monochromatic L2 = color L4 (range L3-L7)

Monochromatic L5 = color L11 (range L9-L13)

Monochromatic L9 = color L13

Evidently, the relative value of sharpness with respect to color was greater for "girl" than for "fruit." An interesting side result is that of two pictures having same L number (for example, L2 or L5), the one with color was judged by some observers to be actually somewhat sharper.

b. Monochromatic Reproductions – Trained Observers

After having participated in several other tests, including the object-identification test, four of the observers were asked to make the same comparisons as in the first

test (a), now, however, with the usual criteria. As expected, every observer judged the monochromatic reproductions to be equivalent to a more blurred color reproduction than previously. The shift varied from one observer to another (the variation depending on his previous amount of 'bias') and was, on the average, one aperture size for "fruit" and two sizes for "girl." Thus, for example, the color equivalence for "fruit" of monochromatic L2 changed from L13 to L21.

c. Spatially Filtered Chrominance

Reproductions of pictures in which the chrominance was spatially lowpass-filtered by aperture sizes 21 and 37 were compared with the standard series. Observers were almost unanimous in judging their equivalents to be just slightly worse than the fundamental code. The equivalence on the average was

$$XY21 = XY37 = L3$$

for both pictures.

d. Chromaticity Quantization

As a preliminary test, four observers were asked to state their preference with respect to each degree of chromaticity quantization as to the size of lowpass-filtering aperture, if any, to which the chrominance was subjected additionally. As expected from the results on spatially filtered chrominance (c), the differences were slight, but the general preference was for the unfiltered version. Only for the roughest quantization of "girl" (quantum size . 08) was the filtered version generally preferred (XY21 and XY37 equally). For the next quantum size (. 056) the opinion was equally divided.

Now each observer was asked to compare the chromaticity-quantized reproductions preferred by him in the preliminary test with the standard series. The resultant equivalences were spread over the range of 2 values in the standard scale for quantum size . 02 and over 3 values for all other quantizations. The average equivalences established are approximately as follows:

"fruit":	quantum size . 02 = L4	"girl":	quantum size . 02 = L3
	quantum size . 04 = L5		quantum size . 04 = L11
	quantum size . 056 = L9		quantum size . 056 = L17
	quantum size . 08 = L12		quantum size . 08 = L21
	quantum size . 1 = L12		

Evidently, the effects of quantization proved to be more objectionable in a picture to which a criterion of an aesthetic nature was applied than where recognizability was paramount.

An additional test was run — to compare chromaticity-quantized reproductions

directly with a monochromatic one. For "fruit" color pictures with up to .056 chromaticity quantum size were preferred to the monochromatic. For "girl" color was preferred only with quantum size .02, while the reverse was true for coarser quantization.

e. Luminance Quantization

A preliminary test established the preference of the observers as to spatial luminance filtering, if any, for each degree of luminance quantization. The majority of observers preferred no filtering (that is, sharp pictures) for "fruit" and "girl" quantized to 18 levels and for "fruit" quantized to 9 levels. The "girl" quantized to 9 levels was preferred by all with a medium-aperture (L13) filtering.

The luminance-quantized and filtered pictures were then compared with the standard series. The results can be stated as follows: Pictures quantized to 18 levels were generally preferred to unquantized ones if the latter were filtered with a larger aperture than the former (for example, 18-level quantized L9 was preferred to nonquantized L13, although, of course, nonquantized L9 was preferred to both). The same was true for the 9-level quantized L13. Nine-level quantized L2, however, was judged to be equivalent, variably, to between L5 and L21 (or even worse for "girl").

An additional test run here was a direct comparison of luminance-quantized color reproductions with the fundamentally coded monochromatic reproduction. Except for the 9-level L2, the color pictures (namely 9-level L13 and 18-level L2 and L9) were generally preferred.

f. Simultaneous Chromaticity and Luminance Quantization

Following the same pattern as in the tests of chromaticity quantization (d) and luminance quantization (e), we sought, first, the effects of spatial filtering of luminance and chrominance on subjects' preferences. Only luminance-dependent chrominance quantization was applied. All significant preferences established here are in line with previous tests in which these effects were studied separately.

On comparison with the standard series, the dependently quantized pictures were judged to have approximately the average equivalences given in Table 2.

Comparison with the monochromatic pictures established preference for color in "fruit" with the first four combinations in Table 2. In "girl," color was preferred with the first and third combination but monochromatic was preferred with the fourth. All other comparisons were inconclusive.

Finally, an attempt was made to rank-order all of the combinations of chrominance and luminance quantization, both independent and dependent, that were carried out. A fairly close consensus of three observers established the descending orders of preference shown in Table 3. The total number of values are the estimated numbers of quantum 'bricks' in the solid of reproducible colors. We find that they are arranged more or less in a descending order, especially for the "fruit." The only exceptions to this order, besides the ones that can be corrected by interchanging two codes ordered successively,

Table 2. Equivalences between luminance quantized and dependently chromaticity quantized color pictures and blurred color pictures.

Chromaticity quantization	Luminance quantization	Filtering	Equivalent unquantized	
			"fruit"	"girl"
fine	36 levels	L2	L4	L5
coarse	36 levels	L2	L4	L13 (wide variance)
fine	18 levels	L2	L4	L8 (wide variance)
coarse	18 levels	XY21L13	L13	L13
rough	9 levels	L2	incomparable	L37?

are for "fruit": .01, 18, having the highest rate, belongs to the beginning of the list; coarse, 36 belongs after .02, 9, while .02, 9 belongs after .04, 36. For "girl" the notable exceptions again involve .02, 9 and coarse, 36. These, as well as the minor anomalies in the order, indicate (i) that luminance-dependent chrominance quantization does not improve quality over independent quantization for same transmission rate, and in some cases it even results in lower quality; and (ii) that luminance quantization, once noticeable, has a worse effect on picture quality than an equivalent (in terms of per sample transmission rate reduction) chrominance quantization.

3.3 SUBJECTIVE RECOGNITION TESTS

Asking an observer to judge the quality of a reproduction by a high-order (in the chain of perceptual decoding) criterion, no matter how well defined, such as 'recognizability of objects', is, of course, no guarantee that this will be the only criterion actually brought to bear. There may, in fact, be two disturbing influences: (i) some undefined criteria of aesthetic nature, involving lower order perception, may affect the over-all judgment, and (ii) it is difficult for the observer to objectively evaluate the recognizability of objects after he has already recognized them in the present or previous tests. A much more objective and reliable method of testing for such a criterion would be simply to ask observers to identify objects in the picture. (It is partly for the straightforwardness and ease of such a test that the criterion of recognizability was chosen for the "fruit" picture in general.)

Because of the limitations on time and on availability of observers, these tests were run only on the spatially filtered color and monochromatic pictures. Of course, only the "fruit" scene was used here. A test was conducted as follows: A group of observers, numbering from 2 to 5, was shown the series of transparencies projected one at a time. They were each asked to write down on the first line of a specially ruled sheet of paper

Table 3. Order of preference for various quantization schemes.

"fruit"			"girl"		
Chromaticity quantization	Luminance quantization	Total number of values	Chromaticity quantization	Luminance quantization	Total number of values
.02	36	4860	.02	36	4860
fine	36	3350	.01	18	9350
.01	18	9350	.02	18	2430
.02	18	2430	fine	36	3350
fine	18	1675	.04	36	1260
.04	36	1260	fine	18	1675
.056	36	575	.04	18	630
.04	18	630	.056	36	575
coarse	36	1040	.08	36	396
.1	36	288	.056	18	288
.056	18	288	.02	9	1215
.02	9	1215	coarse	36	1040
rough	36	432	.08	18	198
rough	9	108	rough	36	432
.08	18	198	rough	9	108
.08	9	99	.08	9	99

the name of every object in the first picture which they could positively identify, then to write on the second line the name of every object that they could identify in the second picture in addition to, or as a correction to, what they had previously identified, and so on for every successive picture. Each picture was shown for a duration of 1-2 minutes, the time depending on the amount of writing that it evoked – apparently enough for most observers to make and write down most of their potential identifications. Each group (except for the smallest ones) was shown the series of pictures in a different order, some with monochromatic first, some with color first – both usually in the order of decreasing aperture size (with some exceptions, intended as a check). There were altogether some 30 observers, including some of those used in the comparison tests, that is, some who had been given the spontaneous judgment test on monochromatic transparencies (a). Every sheet was subsequently checked and the number of correct identifications for each picture noted.

The total number of correct identifications by each observer varied from 9 to 20. A good portion of them was usually made on the first few pictures (as will be detailed below); then additional ones were made as higher and higher quality pictures were shown, up to and including the sharpest color picture L2 (which was normally shown last). The results were summarized in the following manner.

First, the fraction of correct identifications on the picture shown first was determined by summing the numbers of correct identifications on the first picture recorded by all observers who had been shown a particular picture first, then summing the total numbers of correct identifications made by these observers, and finally dividing the first sum by the second. The results are

L9 monochromatic shown first:	61/203 = .30 (13 subjects)
L2 monochromatic shown first:	47/117 = .40 (8 subjects)
L37 color shown first:	84/176 = .48 (11 subjects)
L13 color shown first:	54/74 = .73 (5 subjects)

Second, the fraction of correct additional identifications on the picture shown second was calculated in a similar manner for two combinations of first and second picture. These were

L2 monochromatic shown after L9 monochromatic:	9/71 = .13 (4 subjects)
L37 color shown after L9 monochromatic:	38/119 = .32 (8 subjects)

Third, the fraction of correct identifications made altogether on two pictures shown first was calculated, as above, for various combinations:

L9 monochromatic first, then L2 monochromatic:	29/71 = .41 (4 subjects)
L9 monochromatic first, then L37 color:	76/119 = .64 (8 subjects)
L2 monochromatic first, then L37 color:	41/64 = .64 (5 subjects)
L37 color first, then L21 color:	70/101 = .69 (6 subjects)
L37 color first, then L5 color:	40/55 = .73 (4 subjects)

When similar fractions are computed for individual observers, they vary greatly from one observer to another. For example, for 'L37 color first' it varied from 4/16 = .25 to 15/19 = .79, as extremes. Still, it is felt that the number of observers was sufficiently large to make the results shown above significant for the conclusions to be drawn. One justification for this confidence is that, for each of the first four cases listed, the total number of identifications, divided by the number of observers, was substantially the same; it varied between 14.8 and 16 per observer.

The most striking results appear in the numbers relating to identifications made on transparency viewed first. The number of identifications on L37 color being somewhat larger (though perhaps not significantly) than that on L2 monochromatic implies that L37 color is, for this purpose, at least as good as L2 monochromatic. This is a most important result. The substantially increased number of identifications of L13 color over L37 color verifies the validity of this test by establishing that L37 is, indeed, for this purpose, a degraded version of L13. Similar verification is provided by the quite significant

improvement in L2 monochromatic over L9 monochromatic.

The results relating to identifications made on two successive transparencies are significant only inasmuch as they show a marked advantage of two blurred color versions (L37 followed by L21) over two much sharper monochromatic versions (L9 followed by L2). This concurs with the previous result (namely, since L37 color ~ L2 monochromatic, L37 color must be better than L9 monochromatic and L21 color better than L2 monochromatic). The other results show only that approximately the same fraction of correct identifications is made on the first two transparencies, no matter what they are, if at least one of them is in color; the same fraction holds for L13 color on first viewing. This inconclusiveness brings out possible inadequacies in the test and for the particular picture subject matter used: There may not be enough objects that are unidentifiable in the lower quality versions to leave a sufficient statistically significant number to be identified in subsequently shown versions. Other faults of the test may involve interactions between the successively shown versions, such as failure of observers to overcome misleading clues from a degraded version, and, on the other hand, carrying over of helpful clues and tentative identifications, and the effects of personal factors such as fatigue, interest, and previous experience. Also, the function of the duration of viewing should perhaps have been brought under closer scrutiny. One encouraging result in this context is that whenever a blurred picture was shown after a sharper (but chromatically the same) version, none or an insignificant number of additional identifications was made.

IV. CONCLUSIONS

4.1 DISCUSSION OF RESULTS

Three sets of conclusions may be drawn from my experimental results relating to (a) the minimum parameter values of the fundamental code which allow us to obtain the full quality of reproductions from my experimental system (as a basis of comparison with other experimental and practical systems), (b) coding of chrominance, and (c) color vs monochromatic coding.

a. Parameter Values of the Fundamental Code

Before any code reductions can be discussed we must clearly state our starting point, namely the fundamental code. We must do it for three reasons: (i) Any quantitative values relating to reduced transmission rates are meaningful only when compared with those of the original code. (ii) Even the relative reductions achievable may be quite different if the fundamental code is different. (iii) Comparison of the quality of my variously coded reproductions with that of other experimental and practical transmission systems is meaningful only if the qualities of the respective fundamentally coded pictures is first compared and, preferably, made equal.

The fundamental code is characterized, first of all, by a square array of 256×256 samples. A subjective estimate of the sharpness of the reproductions put the equivalent number of samples at one-half of that, namely at a total of approximately 32,800 samples. The luminance is quantized logarithmically to approximately 36 levels, thereby making it indistinguishable from unquantized luminance. The chrominance is quantized to approximately 1040 values, by squares of size .01 on the Uniform Chromaticity Scale, thereby rendering it indistinguishable from unquantized chrominance. By assuming zero-order probability coding, these parameter values will require a transmission rate of approximately 5 bits per luminance sample and 10 bits per chrominance sample. Thus, our fundamental code requires a transmission rate of 15 bits per sample of $15 \times 32{,}800 = 492{,}000$ bits per picture.

As discussed in Section I, part of the coding process (as far as the observer's eye is concerned) is the manner of reproducing the transmitted picture. Picture reproduction in my system has been described in detail in Section II, and, to be strict, should be construed as part of the fundamental code. Different reproduction (by using, for example, different cathode-ray-tube spot shape, different intensity contrast range, different primary colors or even different film and development) may result in superior or inferior quality. Furthermore, it may result in quantitatively different results, as to the minimum parameter values of the fundamental code, stated above, as well as to the possible reduction thereof. Therefore, the main conclusions reached in this report are essentially of a qualitative nature, and quantities given must be taken to be approximate.

b. Coding of Chrominance

The first and foremost conclusion to be drawn from the results reported in sections 3.1a and 3.2c is that lowpass spatial filtering of chrominance which reduces its effective two-dimensional 'bandwidth' by as much as 18 times does not noticeably affect the quality of picture reproductions such as those made by my experimental system. The factor of 18 is approximately the ratio of the maximum size of convolving aperture, 37 sample points, to the minimum size, effectively 2 sample points. The maximum size was determined by computer memory limitations. It is very likely that even larger factors of reduction of chrominance bandwidth are possible without noticeable effects. The fact that a slight amount of blurring (judged to be equivalent to that produced by filtering luminance with an aperture of size 3) did manifest itself in chrominance-filtered pictures is explainable by the possibility that chrominance variations affect luminance in the actual reproductions.

This conclusion, that chrominance can be greatly blurred (or 'defocused'), was to be expected from a known psychophysical fact, namely that the acuity of the eye to chrominance variations is much smaller than its acuity to luminance variation. Since the highest spatial frequencies of chrominance variation to which the eye is sensitive are lower than those of luminance variation, it is unnecessary (and therefore wasteful of channel capacity) to include any higher frequencies of chrominance in the reproduction. The N.T.S.C. standards for color television specify chrominance bandwidths that are smaller than the luminance bandwidth by a factor in the range from 2.7 for one chromaticity axis to 8 for another axis. These reductions are in the horizontal direction only. If they were extended to include also the vertical direction, these factors would be squared and thus vary from approximately 7 to 64. Our factor of 18 falls within this range.

The results of sections 3.1b and 3.2d show that on the basis of noticeability alone (and therefore for any other possible criterion) chrominance can be quantized into 1040 chromaticities for correct reproduction. This requires transmission of approximately 10 bits per chrominance sample (with uniform probability distribution assumed). If the number of chrominance samples is equal to that of luminance (that is, no spatial filtering), the over-all rate would be 15 bits per sample. The same rate is obtained if we assign five bits to specify the intenstiy sample of each of the three primary colors, which seems reasonable. Therefore, 15 bits per sample was taken to be the transmission rate of the fundamental code. If we now reduce the density of chrominance samples 18 times, as above, it will require an average of 10/18 or 0.55 bits per luminance sample to reproduce the complete chrominance. Thus the over-all average rate per sample is reduced from 5 + 10 = 15 bits to 5 + 0.55 = 5.55 bits.

The results also indicate that any rougher quantization of chrominance than into 1040 values degrades the picture and, moreover, they give the subjectively equivalent degradations if the pictures were blurred by lowpass-filtering both luminance and chrominance.

Let us examine their implications for savings in transmission rate. Consider the apparently least degraded of the noticeably quantized pictures — the "girl" with chrominance quantum size .02. The number of chromaticity values is approximately 270, which corresponds to approximately 8 bits per chrominance sample. Under the assumption of no spatial filtering of chrominance, the over-all rate would be 8 + 5 = 13 bits per sample, a saving of 2 bits from the fundamental 15 bits. The equivalent-quality unquantized picture, on the other hand, whose luminance and chrominance were filtered by an aperture of size 3, requires only 2/3 of the number of samples and therefore an average transmission rate of $2/3 \times 15 = 10$ bits per sample, a saving of 5 bits. All other examples indicate an even smaller relative saving by chrominance quantization. If we now assume spatial filtering of chrominance with aperture of size 37, which was shown to have negligible effect on picture quality, we have an effective transmission rate of 5 + 0.55 = 5.55 bits per sample for unnoticeably quantized pictures. Any rougher quantization could reduce only the 0.55 term (for example, to $8/10 \times 0.55 = 0.44$ for quantum size .02), while spatially filtering the luminance to obtain an equivalent quality would reduce the 5 term (in the example considered, to $2/3 \times 5 = 3.3$). Clearly the savings in transmission rate which are achievable by quantizing the spatially filtered chrominance are very small indeed when the resulting degradation in quality is considered. For the same quality one could code the picture much more efficiently by simply reducing the density of the luminance samples.

It is possible that the number of chrominance quantum values could be reduced without noticeable effects, by using a different geometry of divisions in the chromaticity plane than the equal squares used in my system. Further reduction of this number may be possible if some method is used to decrease the noticeability of spurious contours. One such method would be a modification of the Roberts method for brightness quantization,[17] namely to add pseudo random noise to the X and Y chrominance quantities before quantization and subtract the same noise after reception. Another method would be to simply lowpass the chrominance after reception. This method has been tried, but without much success, as indicated by our results in sections 3.1e and 3.2d. Evidently, the aperture used (size 37), which did not affect natural contours, was also too small to visibly affect the spurious contours. A larger one, on the other hand, may affect both and thus again be impractical. More complicated schemes could perhaps render some improvement here, for example, pre-emphasizing the high spatial frequencies of chrominance before quantization and subsequent lowpass filtering or performing the filtering only in such local areas of the picture in which chrominance steps occur without accompanying luminance steps.

Aside from spurious contours, quantization causes chrominance distortion. In the present system it was noticeable at quantum size .04. It is hard to say how much of such distortion would be tolerated if the contours were completely blurred; this would depend on the criterion used. The number of quantum values for a given distortion could again be somewhat reduced by using a different geometry. Judging from our

transparencies, however, it is improbable that a number less than 32 would be tolerated. This is approximately the same number proposed by Valensi[13] and approaches the number of 23 psychologically distinguishable colors established for distinct light sources.[18] It sets a lower bound of 5 bits on the transmission rate of a chrominance sample.

It should be pointed out again that in view of the much smaller density of chrominance samples that need be sent (as compared with density of luminance samples), any effort at better chrominance quantization will result in a relatively small savings in over-all transmission rate. For example, with the number of chrominance samples reduced 18 times, the lower bound of 5 bits per chrominance sample, cited above, corresponds to an over-all average rate of 5.275 bits per sample – a mere savings of 0.275 bit. If only one-dimensional filtering is used (as in color television at present), the savings may be somewhat more significant, that is, from 5 + 10/4 = 7.5 bits to 5 + 5/4 = 6.25 bits for this example. The savings in chrominance may of course become significant also if some coding method reduces substantially the effective rate of transmission for the luminance.

c. Coding with Color or Monochromatically

In the preceding section it was concluded that a picture can be coded in color with the best quality achievable on our experimental system by using a transmission rate of 5.55 bits per sample. If we code it monochromatically, it can be transmitted at its best quality at a rate of 5 bits per sample. All of our results show unequivocally that the quality of the color-coded picture is superior to that of the monochromatically coded one, by any criterion used. If we now degrade the color picture somewhat so that it can be transmitted at a rate of 5 bits per sample, will it still have a better quality than the monochromatic one and if so, how much further can it be degraded until the qualities of the two pictures are judged equal? On the basis of the results of sections 3.2a and 3.2b and the end of section 3.2f, the answer to the first part of the question is a definite "yes" because in all of the cases reported color pictures of even lower rate than 5 bits were better than the monochromatic one. The second, quantitative, part of the question will be answered in the following discussion.

As pointed out in the preceding section, the chrominance can be coded so that it is transmittable at a rate of 0.55 bit per sample. In order to achieve any substantial reduction in over-all transmission rate we must reduce the contribution to it by the luminance signal from its fundamental value of 5 bits. This can be done by either using any of the more sophisticated coding techniques that have been developed for monochromatic pictures[1-6,17] or simply degrading the luminance and with it the quality of the picture. The latter approach was the one investigated in our experiments. The degradation was the result of reducing either one or both of the basic parameters: number of quantum levels and effective sample density.

Reduction of the effective luminance sampling density, which is equivalent to spatially lowpass-filtering the luminance, was discussed in section 2.2. As was noted, pictures

whose luminance was filtered to various degrees served as a standard of comparison for other reduced-quality pictures. It so happens, as will be shown, that of all combinations of reduced luminance parameters, the ones involving sample density only were the most efficient in terms of reduced transmission rate for a given quality. For this reason the luminance-filtered color transparencies were the ones tested most extensively for comparison with the monochromatic ones. Two major and one minor tests were run.

The first major test is reported in section 3.2a. The indefiniteness of the criterion used for comparison makes the quantitative results here even more approximate than others in this report. The results show, however, that whatever perceptual criterion (as distinguished from a low-level, sensory criterion such as sharpness) the observer chooses to judge by, a color picture is always preferred over a monochromatic one of the same luminance sample density and, furthermore, the luminance sample density of the color picture can be reduced by more than the factor of 5.55/5 (that is, to have a transmission rate less than that of the corresponding monochromatic picture) before the qualities of the two pictures become equal. The amount of reduction seems to depend on the subject matter and on the degree of sharpness with which one begins. There seems to be some relation between the importance of fine detail and the relative value of sharpness (that is, sample density), but we lack sufficient data to infer this relation. It is easy to see, though, how subject matter can effect the relative value of color for identification of objects. In the "fruit" picture there is a large amount of detail consisting partly of various spherical and elongated shapes, where color is obviously of much informational value. In the "girl" on the other hand there are few and easily recognizable objects whose colors are either self-evident (skin, hair, leaves), unimportant (background) or of interest by themselves (color of clothes); only the geranium flowers are easier to recognize with color. Even in this picture, however, color may contribute in other ways, for example, by enhancing distinguishability of various objects and parts.

The second major test was the identification of objects in the "fruit" picture by a large number of observers. The results, reported in section 3.3, are more conclusive than those of any of the comparison tests because they indicate the relative quality of pictures as determined solely and clearly on the basis of the recognizability criterion. The main result was: If a group of observers is shown successively 6-8 versions of the "fruit" picture and asked to identify the objects in it, the ratio of the total number of correct identifications made on viewing the first-shown version to the total number of correct identifications made on viewing the whole series was slightly higher when the first version was L37 color than when it was L2 monochromatic; the ratio was significantly higher than these when the first transparency was L13 color, and significantly lower when the first transparency was L9 monochromatic. Ratios, rather than absolute numbers, were compared because they represented 'normalized' results; this was in actuality unimportant, however, since the denominators were fairly equal.

The conclusion to be drawn from this result is that on the sole basis of ready identifiability of natural objects in a picture, a monochromatic reproduction is approximately as good as a color version thereof in which the luminance (as well as the chrominance) sampling density has been reduced by a factor of 18. If the necessary transmission rate for the monochromatic picture is 5 bits per sample, it will be only 5/18 + 10/18 = 0.83 bit per sample for the color-coded picture – a reduction by a factor of 6.

Like other conclusions in this report, the conclusion drawn above and particularly its quantitative aspect apply, strictly speaking, only to the picture tested and only to the fundamental code used for its reproduction on our system. In other pictures color may contribute less (as in our "girl" picture) or more to identifiability, or the amount of important fine detail may be such that the fundamental sampling rate is too low to be reduced any further. As a demonstration of possibility and as an indication of potentialities, however, the presented conclusion may be quite significant.

The third, minor, test reported in section 3.2b, represents a middle ground between the two tests discussed above, and so do the results. The fact that, in spite of being told to apply the 'recognizability' criterion to the "fruit" picture, observers still placed the equivalence of monochromatic L2 at a sharper color version than the L37 (demonstrated in the identification test) shows that the criterion of sharpness has unavoidably an effect on the observer's judgment. On the other hand, it seems that the observers did learn to apply the recognizability criterion – even to the extent that it apparently affected their 'aesthetic' judgment regarding the "girl" picture and made them equate its monochromatic version to a more blurred color version than in the first comparison test.

Results of degrading the luminance by reducing the number of quantum levels are reported in section 3.2e. They indicate that the effects of reducing the number of levels to one-half that of the fundamental code are less disturbing than those of reducing the sampling density by a factor of 2.5 but that a smaller number of levels causes a greater distortion. This fact by itself is of little value, since halving the number of levels reduces the transmission rate by only one bit per sample (as compared with the 3-bit reduction for the smaller sampling density). More valuable is the result that the effect of coarser quantization is even less noticeable after spatially lowpass-filtering the luminance: reduction of the number of levels by a factor of 4 was tolerated when filtered by a medium-sized aperture.

This result is somewhat misleading because in my experimental system the filtering was in effect carried out on the full density of samples (after quantization) rather than on the reduced density. It is very likely, however, that the same results will hold true also when the actual sampling density is reduced (with the resultant blurring of natural, as well as artificial, contours), thereby allowing the number of quantum levels to be reduced by a factor of 4 without affecting the picture quality appreciably (by any criterion). [This means that if our fundamental sampling were an array of approximately 50 × 50 sample points only, the fundamental quantization would be perhaps to 18 levels,

while 9 levels would be tolerated. It may even prove true for monochromatic pictures that lower sampling density (that is, reduced spatial bandwidth) allows for coarser quantization.] The required transmission rate would then be reduced from 5 bits to 3 bits per actual luminance sample, that is, by a factor of 3/5. Thus a reduced rate resulting from smaller sample density apparently can be reduced further by this factor. Although pictures thus coded were not subjected to the identification test, it can be assumed, on the basis of these results, that the reduction in number of levels would not have diminished their quality in that test. Hence, a revised estimate of the transmission rate required for a color-coded picture for identifiability equal to that of a 5-bit monochromatic picture would be $10/18 + (5/18) \times (3/5) = 0.72$ bit per sample.

Contrary to the figure of 5.55 bits (the rate for the undegraded picture), the last figure of 0.72 bit receives its main contribution (namely 0.55 bit) from the chrominance term. The question therefore arises whether a further reduction of rate would not be possible by allowing some alternative or perhaps additional degradation of the picture through reduction of chrominance parameters, while the effective equivalence with the monochromatic picture is maintained. Unfortunately, the chrominance sample density has reached the lowest value possible in my experimental system, without the fact being established that this was the lower limit for even the fundamental code. Therefore no transparencies are available in which the effect of noticeable chrominance blurring could be tested. Nor was it possible to test how much reduction of chrominance sample rate will remain unnoticeable once the luminance sample rate has been reduced. It is possible, on the one hand, that the former depends entirely on the acuity of the eye to chrominance change, with the result that any further reduction would become noticeable. On the other hand, it is conceivable that the ratio of sample density of luminance to that of chrominance remains the same as in the fundamental code for no noticeable effect with the result that the blurred-color equivalence of monochromatic pictures, established in the identification test, would have a rate of $5.55/18 = 0.31$ bit per sample.

Results of experiments with reduced chrominance quantization combined with luminance reductions are reported in section 3.2f.

Of the three sharp quantized pictures that were judged equal in quality to L4 color, the one quantized to 36 luminance levels and with chrominance quantized, dependently, 'coarse' has the lowest rate, namely approximately 12 bits. This is, however, higher than the corresponding 7.5 bits of L4. The independently luminance- and chrominance-quantized picture judged (according to the last part of section 3.3f) to be equivalent to 'coarse', 36 was .056, 18. The rate for the latter is approximately 9 bits, or if XY37 chrominance filtering is assumed, it is $4 + 5/18 = 4.28$ bits, compared with $5/2 + 10/18 = 3.05$ bits for L4XY37. The conclusion to be drawn from all of this is that for a sharp color picture any degree of noticeable chrominance quantization, with whatever degree of luminance quantization, will distort the reproduction more than luminance blurring that achieves the same rate reduction

The only promising result is the one that shows that an L13 luminance filtered picture

can have its luminance quantized to 18 levels (as concluded earlier) and also have a 'coarse' dependent chrominance quantization, without affecting its quality (even for the "girl"). This quantization requires a rate of approximately 9 bits per actual sample. At least an equal quality should be achievable by independent chromaticity quantization of step size .056, which with 18-level luminance quantization also requires a rate of approximately 9 bits. The corresponding average rate is 4.2/6.5 + 9/18 = 1.15 bits. If this result were applied to the L37 color picture, its rate, which is normally 15/18 = 0.83, would be reduced to 9/18 = 0.50. This is indeed somewhat better than the figure of 0.72 calculated for the case in which luminance quantization alone was applied.

4.2 SUMMARY OF CONCLUSIONS

(i) A color picture, of the highest quality reproducible on my experimental system, can be coded so as to require a rate of transmission at most of 5.55 bits per sample point. This code consists of all of the luminance samples, quantized logarithmically to 36 levels, and at most of 1/18 that number of samples from a two-dimensional spatially filtered version of the chrominance (having a passband area 1/18 that of the luminance), each sample quantized to 1040 chromaticities. By comparison, the best reproducible monochromatic picture can be coded to require a rate of 5 bits per sample point, by using the same code as the one for luminance.

ii) Color pictures, having the contents as used in my experiments, when coded as above but with the density of luminance samples (and, proportionately, the highest spatial frequency of luminance) reduced just enough to bring the required transmission rate to 5 bits per original sample point, are usually judged (by almost any criterion) to be superior to the best monochromatic versions.

(iii) Color pictures as characterized above can be transmitted at maximum rates ranging from 3.05 bits to 1.15 bit per sample point, by further reducing the number of luminance samples and, also, when this number is small enough, halving the number of quantum values of chrominance and of luminance, and still have their quality judged equal to that of the best monochromatic pictures; the rate depends on the picture and on the criterion for comparison of quality, the highest one applying to the "girl" when no particular criterion is specified, and the lowest, to the "fruit" when recognizability of objects was the basis of judgment.

(iv) If the sole purpose of the picture is to convey the presence of certain identifiable natural objects, as in the "fruit" picture, a best monochromatic picture, coded as above and requiring a transmission rate of 5 bits per sample point, will not be any better than a blurred color picture that requires a transmission rate of only 0.50 (or possibly less) bit per original sample point by coding it as follows: The number of luminance samples is 1/18 of the original number of sample points, the number of chrominance samples is also 1/18 (or possibly less) of that number, luminance is quantized to 18 levels and chrominance to 32 values.

Acknowledgment

My thanks are due first and foremost to Professor William F. Schreiber who has guided me with ideas, practical suggestions and, above all, encouragement and inspiration throughout this work.

I also wish to thank the other members of my thesis committee, Professor Murray Eden and Professor Samuel J. Mason, for their helpfulness, and James E. Cunningham, John W. Pan, and Oleh J. Tretiak, members of the Picture Processing Group of the Research Laboratory of Electronics, for their valuable cooperation in building and operating the recording-reproducing equipment.

The contribution of Carolyn S. Shane, Rosina C. Petone, Donald N. Graham, Ralph E. Grabowski, Thomas S. Huang and, again, Oleh J. Tretiak, who served as observers in the evaluation of the picture reproductions, is also gratefully acknowledged.

This work was carried out at the Research Laboratory of Electronics, Massachusetts Institute of Technology, whose support I gratefully acknowledge. The computer processing was carried out at the Computation Center, M.I.T.

References

1. E. R. Kretzmer, Reduced Alphabet Representation of TV Signals, IRE Convention Record, Part IV, pp. 140-147, 1956.
2. W. F. Schreiber and C. F. Knapp, TV Bandwidth Reduction by Digital Coding, IRE Convention Record, Part IV, pp. 88-98, 1958.
3. W. A. Youngblood, Estimation of Channel Capacity Required for Picture Transmission, Sc.D. Thesis, Department of Electrical Engineering, M.I.T., June 1958.
4. M. Corradetti, A Method of Planar Prediction for Coding Pictures, S.M. Thesis, Department of Electrical Engineering, M.I.T., July 1959.
5. B. Julesz, A method of coding TV signal based on edge detection, Bell System Tech. J. 38, 1001-1020 (1959).
6. J. W. Pan, Reduction of Information Redundancy in Pictures, Sc.D. Thesis, Department of Electrical Engineering, M.I.T., September 1962.
7. D. Fink (ed.), N.T.S.C. Standards on Television (McGraw-Hill Book Company, Inc., New York, 1943).
8. D. Fink (ed.), N.T.S.C. Standards on Color Television (McGraw-Hill Book Company, Inc., New York, 1955).
9. E. C. Cherry and G. G. Gouriet, Some possibilities for the compression of television signals by recoding, Proc. I.E.E. (London), Vol. 100, Part III, pp. 9-18, January 1953.
10. G. G. Gouriet, Bandwidth compression of a television signal, Proc. I.E.E. (London) 104B, 265-272 (1957).
11. Committee on Colorimetry (L. A. Jones, Chairman), Opt. Soc. Am., The Science of Color (Thomas Y. Crowell Company, New York, 1953).
12. A. C. Hardy, Handbook of Colorimetry (The Technology Press of Massachusetts Institute of Technology, Cambridge, Mass., 1936).
13. G. Valensi, Note on a method of coded colour television, Communication Theory, edited by W. Jackson (Academic Press, New York, 1953), pp. 354-359.
14. Picture Processing Group, Research Laboratory of Electronics, M.I.T., unpublished notes.
15. F. C. Breckenridge and W. R. Schaub, Rectangular uniform-chromaticity-scale coordinates, J. Opt. Soc. Am. 29, 370 (1939).
16. O. J. Tretiak, The Picture Sampling Process, Sc.D. Thesis, Department of Electrical Engineering, M.I.T., June 1963.
17. L. G. Roberts, PCM Television Bandwidth Reduction Using Pseudo-Random Noise, S.M. Thesis, Department of Electrical Engineering, M.I.T., February 1961.
18. K. L. Kelly, Color designations for lights, J. Opt. Soc. Am. 33, 627 (1943).